REMARKS

Claims 119, 122 and 123 have been amended, and new claims 127-128 have been added. Thus, claims 119-128 are pending. No new matter has been added. In view of the above amendments and the remarks below, Applicants respectfully request that all of the pending claims be allowed.

Claims 119-126 have been objected to for containing language directed to intended use and for formatting issues. In view of the above amendments to claim 119 and 123, Applicants respectfully request that these objections be withdrawn.

Claims 119-126 stand rejected under 35 U.S.C. §112 for failing to particularly point out and distinctly claim the subject matter which Applicants regard as the invention. In view of the above amendments to claims 119, 122 and 123, Applicants respectfully submit that this rejection has been overcome.

Claims 119-126 stand rejected under 35 U.S.C. §101 for being directed to non-statutory subject matter. In view of the above amendment to claim 119, Applicants respectfully submit that this rejection has been overcome.

Claims 119-120 stand rejected under 35 U.S.C. § 103(a) as unpatentable over Fukuda, Kenichi, JP-08063483 ("Fukuda") in view of Isozaki, Hidaki, JP-2001-318792A ("Isozaki").

Claim 119 recites a method for automating the extraction of information from a semi-structured document characterized by a document type that comprises design and structural characteristics of a set of similar documents comprising "designing a target extraction template for terms of the document type" and "supporting the creation of a control set of documents containing terms manually tagged to the extraction template" in addition to "automatically generating a skeleton of an extraction model tree for every term" and "identifying a set of

selectors for each model tree" and "training the model trees by automatically identifying a subset of the selectors for the extraction model trees for compliance with the control set" and "extracting information from the document with the optimized model trees" and "storing the extracted information in a database."

In contrast, Fukuda discusses a system for converting text data within a document structure into card type data with standard columns for output of such text data. Fukuda ¶12. The Fukuda system checks the entirety of the text data for reserved words and, for each reserved word located, creates card type data with the text data that corresponds to the reserved word. *Id.* at 21-23. Isozaki discusses a system for generating intrinsic representation extraction rules and selecting those rules which properly classify text within documents. Isozaki, ¶¶9-10.

Neither Fukuda nor Isozaki, either alone or in combination, discloses or suggests "identifying a set of selectors for each model tree" and "training the model trees by automatically identifying a subset of the selectors for the extraction model trees for compliance with the control set," as recited in claim 119. The Examiner acknowledges that Fukuda neither discloses nor suggests model trees but states that Isozaki discloses the claimed model trees and discloses, at ¶¶49-50, training the models trees by automatically optimizing selectors for the model trees.

In Isozaki, the value 't' (for type of intrinsic representation), '+df' (for character-shifting to the right from the start position of the intrinsic representation, and '-dt' (for character-shifting to the left from the end position of the intrinsic representation) are used to construct an empirical rule for identifying the intrinsic representations within a document. *Id.* at ¶¶47-49. In contrast to the present invention, Isozaki uses all of these values to create each of the empirical rules. *Id.* at ¶58. Thus, Isozaki neither discloses nor suggests, "training the model trees by automatically identifying a subset of the selectors for the extraction model trees for compliance with the

control set," as recited in claim 119.

Therefore, it is respectfully submitted that neither Fukuda nor Isozaki, either alone or in combination, discloses or suggests "training the model trees by automatically identifying a subset of the selectors for the extraction model trees for compliance with the control set."

It is respectfully submitted that Bernstein et al., "Discovering Knowledge from Relational Data Extraction from Business News," Stern School of Business, New York University, NY, CeDER Working Paper #IS-02-03 ("Bernstein") does not cure the deficiencies of Fukuda and Isozaki. That is, Bernstein discusses methods of discovering relationships between documents (and between companies identified in those documents) by analyzing the frequency of words occurring in the documents. Bernstein never discloses or suggests "training the model trees by automatically identifying a subset of the selectors for the extraction model trees for compliance with the control set," as recited in claim 119.

Therefore, it is respectfully submitted that neither Fukuda nor Isozaki nor Bernstein, either alone or in combination, discloses or suggests "training the model trees by automatically identifying a subset of the selectors for the extraction model trees for compliance with the control set," as recited in claim 119. Accordingly, Applicants respectfully submit that claim 119 is allowable, and the claims depending therefrom (120-126) are also allowable.

Claim 127 recites a method comprising "identifying a plurality of indicators in a document type" and "generating a decision tree for the document type based on a subset of the plurality of indicators" in addition to "identifying a location of a term within the document type as a function of the decision tree" and "comparing the location of the term with a control location for the term in the document type" and "generating an extraction template for the document type."

For at least the reasons stated above, it is respectfully submitted that neither Fukuda nor Isozaki nor Bernstein, either alone or in combination, discloses or suggests "generating a decision tree for the document type based on a subset of the plurality of indicators," as recited in claim 127. Accordingly, Applicants respectfully submit that claim 127 is allowable, and the claim depending therefrom (128) is also allowable

Please deduct any fees due from Account No. 50-4798.

Dated: April 23, 2009   Respectfully submitted,

                                                          By: _____
                                                          Timothy M. Salmon
                                                          Registration No. 54,420

                                                          ALBERT JACOBS LLP
                                                          Customer Number: 84310
                                                          900 Third Avenue, Suite 502
                                                          New York, New York 10022
                                                          Tel: (212) 508-7533
                                                          Fax: (212) 508-7541